

学校编码: 10384

学号: 31920131153038

分类号____密级____

UDC____

厦 门 大 学

硕 士 学 位 论 文

网络负面新闻信息编码分类与检索关键字研究

Internet Negative News Coding for Classification and
Key Word for Retrieval Research

杨旭彬

指导教师姓名: 黄合水教授

专 业 名 称: 广告学

论文提交日期: 2016年3月

论文答辩时间: 2016年5月

学位授予日期: 2016年 月

答辩委员会主席: 曹立新

评 阅 人: 蔡卫平 查本恩

2016年5月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名): 杨旭彬

2016年5月20日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：杨旭彬

2016年5月20日

摘 要

与区域相关的网络负面新闻信息逐渐成为了区域内执政者维护区域媒介形象和危机管理的焦点，因此对负面新闻信息的分类和检索就显得意义重大，而开发高质量的负面新闻信息编码分类工具和检索工具则是进行负面新闻信息分类和检索的关键。为解决负面新闻信息分类与检索这两个问题，笔者做了两项研究：研究一是运用扎根理论，对混合抽样所得的 6626 条负面新闻标题进行了归纳，力图得到涉及区域媒介形象的网络负面新闻信息编码分类工具、和分类类目；研究二是对 6626 条负面新闻信息进行关键字词的归纳，力图得到区域媒介形象各主类目下的负面新闻信息的检索关键字词。研究结果如下：

(1) 负面新闻信息编码分类工具，由腐败、社会治安、自然灾害、意外事故、民生问题、环境破坏和经济发展 7 个核心类目，及其所包含的 24 个主类目构成，7 个核心类目的信度检验值均在 0.9 以上，各核心类目和主类目得到了编码时详细的操作定义；成功分类后的负面新闻信息类别中，不同核心类目和主类目所占比例不同，表明媒介通过负面新闻信息塑造区域媒介形象时也有不同的侧重点。

(2) 负面新闻信息的检索关键字，由 24 个主类目的关键字词构成。其中有 2 个主类别即其他类社会治安问题、和黄赌毒类社会治安问题，的关键字词只有一组，如黄赌毒类社会治安问题用“毒品、卖淫”等词就可检索出该类别下的负面新闻信息。有 1 个主类别即破坏植被与伤害野生动物，的关键字词需由两组检索，它们是“收售、盗伐”等词，和“熊、树”等词；其余 21 个主类目的关键字词均由三组构成，如生活腐败主类别的第一组中含有“生活作风腐化”等词，就可检索出该类别下的负面新闻信息，而剩余两组需交叉组合才可检索出该类别下的负面新闻信息。

采用研究所的关键字词进行检索，98.7%的负面新闻信息可以被检索出来，关键字检索出的信息中非负面新闻信息最低容量为 13.9%，最高为 84.7%，如果用这些关键字协助人工编码，平均工作效率可提升 46.8 倍。

关键词：媒介形象；负面新闻信息；编码分类；检索关键字

Abstract

The regional negative news information are so important for the government to manage regional media image and crisis that it is significant to classify and retrieve the negative news information. Therefore developing the tools of classifying and retrieving the negative news information is the key. To get the tools, researcher did two studies. In the first study, researcher generalized the number of 6626 mixed sampling negative news headlines by the means of Grounded Theory in order to get the coding tool for classifying the negative news information. In the second study, researcher also generalized the number of 6626 mixed sampling negative news headlines in order to get the key word for retrieving the negative news information. So the results are as follows.

The coding tool for classifying the negative news information is constituted by 7 core categories that including 24 main categories, which are corruption, social security issues, natural disaster, accident, livelihood issues, environmental damage, economic development issues. And the reliability of each category is above 0.9. Also every core category got the detailed operational definition. Meanwhile, different core categories and different main categories had different proportions that indicates the media had their focus when they shaped the regional media image by negative news information.

The key word tool for retrieving the negative news information is constituted by groups key words of 24 main categories. First, Other social security issues, pornography-gambling-drug social security issues two main categories' key words only have one group. For example, using the words of drug, prostitution and gambling can retrieve the negative news information that belongs to pornography-gambling-drug social security issues. Second, damaging plants and the wildlife one main category's key words has two groups, which are the words of unlawfully fell, purchase and bear, tree and so on. Third, the remaining 21 main categories' key words have three groups. For example, using the word of life style corruption in life corruption main category can retrieve the negative news information

that belongs to this category. And using the words of director, the mayor and commit adultery, mistress together can retrieve the negative news information that belongs to this category.

When using these key word for retrieving, you will get 98.7% of the negative news information. Non-negative news information account for lowest 13.9% of the retrieved news information, and highest 84.7% of the retrieved news information. Also if you use these key word to retrieve and code, the average work efficiency can improve 46.8 times.

Key Word: Media Image; Negative News Information; Coding for Classification; Key Word for Retrieval

目 录

1. 绪论.....	1
1.1 新闻信息类数据特点.....	1
1.2 研究意义.....	1
2. 文献综述.....	3
2.1 负面新闻信息研究.....	3
2.1.1 负面新闻信息概念.....	3
2.1.2 负面新闻信息的负面作用.....	3
2.2 区域媒介形象研究.....	4
2.2.1 媒介形象理论来源.....	4
2.2.2 区域媒介形象概念.....	6
2.2.3 区域媒介形象相关研究.....	8
2.3 研究问题.....	10
3. 研究一：编码分类.....	12
3.1 研究方法.....	13
3.1.1 样本采集.....	13
3.1.2 分析步骤.....	15
3.2 研究结果.....	20
3.2.1 编码分类工具.....	20
3.2.2 信度验证.....	25
3.2.3 不同类别占比.....	26
3.3 研究小结.....	27
4. 研究二：检索关键字.....	28
4.1 研究方法.....	28
4.1.1 样本.....	28
4.1.2 分析步骤.....	28
4.2 研究结果.....	29
4.2.1 检索工具.....	29
4.2.2 效度验证.....	41
4.3 研究小结.....	44
5. 综合讨论.....	45

5.1 研究总结	45
5.2 研究不足	46
参考文献	47

厦门大学博硕士论文摘要库

Contents

1.Introduction	1
1.1 Characteristics of News Data	1
1.2 Research Meanings	1
2. Literature review	3
2.1 Studies of Negative News	3
2.1.1 Concepts of Negative News	3
2.1.2 Negative Effects of Negative News	3
2.2 Studies of Regional Media Image	4
2.2.1 Theories of Media Image	4
2.2.2 Concepts of Regional Media Image	6
2.2.3 Existing Studies about Regional Media Image	8
2.3 Research Questions	10
3.Study One: Coding Tool For Classification	12
3.1 Methodology	13
3.1.1 Sampling	13
3.1.2 Progress	15
3.2 Results	20
3.2.1 Final Coding Tool For Classification	20
3.2.2 Reliability of the Coding Tool	25
3.2.3 Proportion of Different Classifications	26
3.3 Study One Summary	27
4.Study two: Key Word for Retrieval	28
4.1 Methodology	28
4.1.1 Sampling	28
4.1.2 Progress	28
4.2 Results	29
4.2.1 Fianl Key Word for Retrieval	29
4.2.2 Validity of the Coding Tool	41
4.3 Study Two Summary	44
5.Dicussion	45

5.1 Summary	45
5.2 limitations.....	46
Reference.....	47

厦门大学博硕士论文摘要库

1. 绪论

1.1 新闻信息类数据特点

大数据这一兴起于上世纪 80 年代的概念（喻国明等，2013），随着理论界和实务界的不断研究、探讨和实践，使得这一发端于计算机领域的研究已经延伸到其他领域和学科，并产生了深远的影响。大数据具有“数据量大、类型繁多（结构化和非结构化）、难挖掘、变化快等特点（喻国明等，2013；刘春成，2014；樊伟红等，2012；陶雪娇，2013；韩晶，2013）

陈力丹(2015)在讨论大数据时认为，大数据是指在互联网中保留下来的远超出传统数据库的海量数据，而每年在互联网上产生的数以百万，甚至千万计的新闻信息作为人类在使用互联网时所留下的重要痕迹之一，也是大数据构成的重要组成部分。因而网络新闻信息也就相应的有了大数据所有的“规模巨大、非结构化和难以挖掘检索、变化快”等特点。

1.2 研究意义

在如今的大众传播媒介迅猛发展的年代，媒介对人们的影响不言而喻。在我国，截止到 2015 年 12 月，网络新闻拥有用户规模为 5.64 亿人，占全国网民的 82%，是继即时通信、搜索引擎之后使用率最高的网络应用（CNNIC，2016）。这意味着我国有 5.64 亿的用户接受来自网络媒介的新闻信息，接受网络媒介所建构的媒介现实，被网络媒介所影响。而铺天盖地的网络负面新闻信息也恰好隐藏在其中。

一个品牌的长期成功依赖于品牌在消费者心目中保持良好的形象，为了维护企业声誉，企业除了要保持产品品质稳定、让消费者在品牌接触点上都感到满意，更需要及时处理消费者对品牌的负面评价（冯海峰，2013）。一个品牌尚且如此，对于一个行政区域而言，更是如此。张恩韶（2011）在研究网络舆论危机下中国政府的形象塑造时发现：网络舆论危机是政府在互联网发展迅猛阶段遇到的一个重大挑战，政府的负面信息经由网络即时、大范围的传播，易引起社会公众对政府权威的质疑和挑战。因此在互联网迅猛发展，我国网络新闻受众已经达到 5.64 亿，并且还在日益增长的年代，如果一个区域想拥有一个良好的区域媒介形象，

那么区域内网络负面新闻信息是一个区域内执政者必须重视的问题。

那么如何将大量的、非结构化的涉及区域媒介形象的网络负面新闻信息类数据结构化, 将此类负面新闻信息成功分类, 了解其分布状况? 同时如何在浩瀚如海洋的、非结构化的负面新闻信息类数据中有效、快捷地检索出不同类别的负面新闻信息? 对于一个亟待掌握地区媒介形象建设主动权的执政者而言至关重要, 通过对网络负面新闻信息分类与检索, 把握自身区域负面新闻信息的布局与走势, 对其了解区域负面舆情走势, 调整区域内信息管理模式, 建设良好的区域媒介形象, 有非常重要的现实指导意义。

2. 文献综述

2.1 负面新闻信息研究

2.1.1 负面新闻信息概念

对于负面新闻信息概念的理解，需要从语义学和新闻学两个角度来理解。

邓利平（2001：5）在其著作《负面新闻信息传播的多维视野》中认为：“负面新闻信息”是一个定中结构的组合词，其中“新闻信息”是中心词，而“负面”是定语，起修饰和限制作用。语义学上，“负面”是个贬义词，有消极、不好的有害的意思，那么直观的从语义学角度来理解的话，负面新闻信息就是不好的有害的新闻信息。那么什么样的新闻信息能够算做是不好的、有害的新闻信息呢？

新闻信息源于新闻事实，而新闻事实则是对现实世界中事物存在、发展、变化的展现，当事物存在、发展、变化有一定新闻价值的时候，新闻信息则会反映这一事实（何淑群，2008）。因此负面新闻信息所反映和报道的是不以人的意志为转移的、不可抗拒的事实（杨保军，1999；邓利平，2001），当新闻信息所报道的事实与人们发生一定联系时，人们就会对这部分事实的性质作出正负之分，而负面新闻信息显然指那些打破人类社会正常的生活秩序，压抑人的自由发展，使得人与人、人与社会、人与自然的关系发生冲突、失衡的事实，例如自然灾害、交通事故、以及现实中的虚假、贪婪、腐败、荒淫、罪恶等等不合乎人类发展的规律性、目的性的活动事实，它们都体现为反常、对抗、冲突。这些新闻信息所报道的消极的、病态的事实对社会的负面影响极大，威胁社会安全，阻碍社会进步（杨保军，1999；何淑群，2008；丁柏铨、陈相雨，2010）。

因此综合上述学者的观点，研究认为负面新闻信息是指具有一定新闻价值的使得人与人、人与社会、人与自然的关系发生冲突、失衡的事实所释放出的信息，从内涵上看，负面新闻信息是指报道的与人类社会发展、与自然发展有悖的事实，从外部表现形式来看，它包含的比较广泛，可以包含各种自然灾害和意外事故，也可以包括违法乱纪、有违道德规范等事实。

2.1.2 负面新闻信息的负面作用

梅尔文·L·德弗勒（1989：370-372）认为媒介可以通过描述和说明的方式，

提出对于现实的解释,这种解释对受众产生潜移默化的影响。而张国良(2001:62)则更深一步的阐明了此类影响,即人们可以从听到的、看到的媒介内容发展出对现实世界的主观和公认的意义构想,随着受众对媒介依赖程度的加深,媒介所提供的信息以及对这些信息的解释,会影响人们对事物的认知、判断,并使得人们据此来决策自己的行动。

从上述两位学者的研究中,反映了两层意思:第一,媒介通过描述和说明的方式刻画被报道者的媒介形象;第二,经过媒介刻画的被报道者形象转而影响受众的认知。负面新闻信息作为媒介刻画被报道者形象的手段和方法之一,会对受众产生影响,也会对被报道者形象产生影响,具体负面影响主要体现在:

第一,负面新闻信息构建负面区域媒介形象。胡秋含(2013)通过研究媒体对于河南负面新闻的报道,发现了媒体通过负面新闻信息所塑造了河南省的负面媒介形象。而其他学者也在研究中,发现媒介运用负面新闻信息所构建的城管、教师、女大学生等人群的负面媒介形象。这也证实了媒介在塑造被报道者媒介形象上有着巨大的影响力(叶兵、蒋兆雷 2007;陶欣龙, 2013;张洋, 2015)。

第二,负面新闻信息对受众认知的危害。邓利平(2001:34-40)谈到负面新闻信息的负面作用时认为,负面新闻信息能够使人精神涣散、破坏人们对于事物的认识,起到一定的破坏式教唆和煽情作用。杨洗和陈怀林(2005)则通过实证研究的方式证明了媒介信息对本地市民城市形象观念的形成具有重要的影响,以上学者均通过各自研究证实了媒介所发出的信息对于受众认知有着巨大的影响。

2.2 区域媒介形象研究

目前学术界所讨论和研究的媒介形象分为两类(栾轶玫, 2007: 17):第一类是关于传播媒介自身形象的讨论与研究,也称传播者媒介形象,第二类是关于事物、个人、组织或地区在媒介中的形象,也称被传播者媒介形象。而本研究将只关注第二类情况下的媒介形象。

2.2.1 媒介形象理论来源

媒介形象作为现实与受众认知之间的桥梁,它的核心在于剖析现实、媒介和认知三者之间的关系。而最早对于这一理念进行阐述的是李普曼1922年《舆论学》一书中所阐述的“拟态环境假说”理论(张国良, 2001: 80)。拟态环境假说理

论认为：大众媒介在反映现实世界时并不是简单重复的再现,而是对事实进行选择、重构再加工后加以传播的,媒介所营造的只是一种“拟态环境”,但受众却以为媒介反映的就是真实的、客观得到现实环境(李普曼,1984:1-23)。随后李普曼(2006:9-12)又在其著作《公众舆论》一书中对这一概念进行了有趣生动进行了解释,大致意思如下:大众媒介通报道为受众构建了一个与客观现实有别的虚拟的现实环境,而这个虚拟的现实环境(即拟态环境)正是媒介受众了解现实、做出判断的参考和依据。而这种具有先天缺陷的参考和依据会严重影响到媒介受众的认识,进而影响到媒介受众的行动,但是在媒介迅猛发展,现实环境庞大、复杂和短暂的情况下,大众又无法不依赖于媒介了解现实、做出判断(韦金艳,2008;王鹏进,2010)。李普曼在“拟态环境假说”理论和随后的著作中,解释了“客观现实、媒介塑造、及受众认知”之间的关系,开创了此类研究的先河。

李普曼的拟态环境假说理论中蕴含了些许“媒介设置议程影响受众”的观点。在1972年,学者麦克姆斯(郭庆光,1999:214)在前人的基础上提出了“媒介议程设置理论”,即大众传媒在塑造社会现实的过程中扮演着重要的角色,具有构建公众思想和引发他们认知变化的能力,它并不能告诉我们应该思考什么,但是可以安排我们去思考什么(李特约翰,2011:340-343)。简而言之,媒介议程设置就是媒介在公众的头脑中建立起突出事务或形象。“媒介议程设置理论”的第一层是寻找具有重要意义的总体性事务,第二层决定这些事务中的哪些方面更为重要,利用一个报道或者传播框架来突出事务的重点,第二层含义中具有明显的“媒介框架”含义(夏倩芳、张明新,2007)。

吉特林(Gitlin)1980年首先把“框架理论”的概念引入到大众传播学研究之中,他利用媒介框架理论研究分析了20世纪60年代哥伦比亚广播公司如何用新闻报道把时兴的学生运动变成了一件无关紧要的事情,随后越来越多的学者开始使用“媒介框架”理论进行研究(Gitlin,1980;Gamson,W.A,1989;McCombs & Shaw,1993;Kanjirathinkal.M.&Hickey J.V.,1992),他们认为媒介框架是媒介对新闻事件的报道的一种框架,可以从新闻文本具有的特征体现出来,通过这种框架限制了受众对事件的理解,而也正是源于这种框架的存在,使得“媒介议程设置”中的第二层作用才能自认而然的产生出来(李特约翰,2011:340-343)。

上述学者的拟态环境假说、媒介的议程设置、媒介框架理论理论,都在强调

媒介塑造对于受众的影响。从他们的学术研究中，我们可以总结出：①媒介是连接受众与社会其他要素之间的重要桥梁；②媒介主观设置的议程、构建的意义和框架对于受众理解世界、做出判断和行动有重要影响；③因为媒介设置的议程、构建的意义和框架中的主观性等原因，导致媒介呈现给受众的意义并不能客观、真实的反映外部世界，因此会对受众理解世界、做出判断和行动产生不利影响。正是基于前人对于理论的不断研究和积累，为接下来“媒介形象”的研究提供了坚实的理论基础。

2.2.2 区域媒介形象概念

对于媒介形象以及区域媒介形象概念的阐述，不同学者均有不同的观点和侧重点，使用和引用较多的定义可以分为以下两大类：

第一类：媒介创造的主观性、受众认知的主观性以及被传者的存在客观性三类因素共同影响媒介形象

张晓莺（2008）认为媒介形象是某个客观事物经运动变化而产生的现实形象或活动状态，经由大众传媒的报道在大众传媒中产生的投影，媒介形象是一个主客观结合的产物，主观性指媒介创造的主观性和受众认知的主观性，客观性指被传者的存在性。而丁柏铨（2009）在研究政府媒介形象的时候所持观点内涵基本与张晓莺相同，即政府的媒介形象是通过大众传播媒介所展示的形象，由政府的现实形象转化而来，因此政府的现实形象是媒介形象的基础和依据，而政府的媒介形象展现与政府自身现实形象、大众传播媒介的报道方式和传播方式、社会公众的接受与认同都有关系。宣宝剑（2008）在其媒介形象系统论中阐述被传者媒介形象时，认为被传者媒介形象是指人们对于大众传播媒介或组织再现的人或事物认知信息的总和，是一种再现形象的认知信息，“被传者媒介形象”既有主观性也有客观性，主观性是指媒介创造的主观性和受众的认知主观性，客观性是指被传者的客观存在及被传者形象信息一经生产也是一种客观存在。此类学者都认为媒介形象具有主观性和客观性，主观性指受媒介创造主观性、受众认知主观性，客观性指被传者的客观存在性及媒介形象信息的客观存在性。

第二类：媒介创造主观性和被传者存在客观性两类因素共同影响媒介形象

江根源等（2006）在研究区域媒介形象时认为，区域媒介形象受传统、权威和刻板映像等主客观条件的影响，会对区域受众认知产生一定的影响。其中主观

Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”.

Fulltexts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.